

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT) (51) International Patent Classification 6: (11) International Publication Number: WO 98/40520 A1 C12Q 1/68, H01J 49/00 (43) International Publication Date: 17 September 1998 (17.09.98) (21) International Application Number: PCT/US98/04919 (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, (22) International Filing Date: 12 March 1998 (12.03.98) GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, (30) Priority Data: KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, 60/040,717 14 March 1997 (14.03.97) US CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, (71) Applicant: HYBRIDON, INC. [US/US]; 620 Memorial Drive, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Cambridge, MA 02139 (US). (72) Inventor: WANG, Bing, H.; Apartment 2, 238 Walden Street, **Published** Cambridge, MA 02140 (US). With international search report. Before the expiration of the time limit for amending the (74) Agents: KEOWN, Wayne, A. et al.; Hale and Dorr LLP, 60 State Street, Boston, MA 02109 (US). claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD FOR SEQUENCING OF MODIFIED NUCLEIC ACIDS USING ELECTROSPRAY IONIZATION-FOURIER TRANSFORM MASS SPECTROMETRY

#### (57) Abstract

The invention provides an analytical method for determining the nucleotide sequence of nucleic acid analytes, including chemically modified oligonucleotides. This new method utilizes electrospray ionization—Fourier transform mass spectrometry.

### FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Моласо	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
СН	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		
							•

# METHOD FOR SEQUENCING OF MODIFIED NUCLEIC ACIDS USING ELECTROSPRAY IONIZATION FOURIER TRANSFORM MASS SPECTROMETRY

### Field of the Invention

The invention relates to the determination of nucleotide sequences for nucleic acids and their analogs.

### Brief Summary of the Related Art

Determination of nucleotide sequences for various nucleic acids has become a fundamentally important analytical step for numerous molecular biology and biomedical applications. Consequently, a variety of methods have been developed to facilitate such nucleotide sequence determinations. Maxam and Gilbert, Proc. Natl. Acad. Sci. USA 74: 5460 (1977) discloses a chemical degradation approach for DNA sequence determination. Sanger et al., Proc. Natl. Acad. Sci. USA 74: 5463 (1977), discloses a chain termination method using complementary strand primer extension to determine DNA sequences. Each of these methods utilizes four separate reaction mixtures to create a nested set of fragments differing by a single nucleotide in length and representing a complete nucleotide sequence, followed by resolution of the fragments based on their size to determine the order of the fragments and hence the nucleotide sequence. Both of these procedures take from numerous hours to days to perform, and neither is effective for determining the nucleotide sequence for certain analogs of DNA.

These popular approaches to nucleotide sequence determination are limited by their time consuming methodologies and by their inapplicability to certain types of nucleic acid analytes. For example, small synthetic oligonucleotides have recently become of interest as tools in molecular biology experiments, as well as for use in the antisense therapeutic approach to disease treatment. Correct sequences are necessary to the efficacy and safety of such oligonucleotides, and

1

effective and rapid analytical approaches are needed for quality control. This class of compounds presents three special problems for traditional sequence determination approaches. First, quality control procedures are needed which are more rapid than the traditional approaches. Second, the oligonucleotides are generally short, often in the range of from about 15 to about 35 nucleotides in length. As a consequence, traditional sequencing approaches result in the loss of much sequence information and thus provide only incomplete information about the nucleotide sequence of the oligonucleotides. Finally, many of these oligonucleotides have either modified internucleoside linkages or substitution at the 2' position of the ribose to improve their properties as potential therapeutic agents. U.S. Patent No. 5,220,007 discloses chimeric oligonucleotides having regions of oligonucleoside phosphodiester or phosphorothioate alongside regions of oligonucleoside alkylphosphonate or phosphoramidate. PCT publication WO94/02498 discloses hybrid oligonucleotides having DNA regions alongside 2'substituted RNA regions. Uhlman and Peyman, Chemical Reviews 90: 544 (1990), discloses oligonucleotides having a variety of modifications along the internucleoside linkages, sugar residues or nucleoside bases.

Recently, some approaches have been developed to address these special problems. U.S. Patent No. 5,403,709 discloses a method for sequencing oligonucleotides using another oligonucleotide as an extension and a third, bridging oligonucleotide to hold the first two together for ligation. Conventional primer extension is then used to create a complement for sequencing. This approach requires some advance knowledge of a portion of the sequence of the analyte oligonucleotide. U.S. Patent No. 5,525,470 discloses a similar approach which avoids the need for such advance knowledge by utilizing RNA ligase to couple the analyte and extension oligonucleotides. Both of these approaches can be used with some modified oligonucleotides, especially with those having phosphodiester or phosphorothioate internucleotide linkages. Both, however, are time-consuming and limited to analyte oligonucleotides which can act as templates for the polymerase enzyme used to synthesize their complement.

Recently, Nordhoff et al., J. Mass Spec. 30: 99 (1995) disclosed a direct sequencing method for oligonucleotides using infrared matrix-assisted laser desorption/ionization mass spectrometry. This method utilized time-of-flight analysis of prompt fragment ions generated by infrared laser desorption/ionization. Fragments were reported to be generated by consecutive cleavage of the deoxyribose phosphate backbone at apurinic sites. Unfortunately, such cleavage does not occur at T residues, due to failure to remove the thymidine base, and may not occur in RNA due to stabilization of the glycosidic bond. In addition, there is a problem with the prompt fragment ions overlapping with fragment ions (PSD ions) from metastable decays occurring in the field-free region of the time-of-flight analyzer.

Brown and Lennon, Anal. Chem. <u>67</u>: 3990 (1995) discloses sequence-specific fragmentation of matrix-assisted laser-desorbed protein/peptide ions and detection of the fragments using time-of-flight mass spectrometry with delayed pulsed ion extraction. The delayed pulsed ion extraction is used to reduce the generation of PSD ions by expanding the desorbed neutral plume during the extraction delay period, thereby avoiding energetic collisions believed to play a role in the generation of PSD. The technique was found to be applicable to small peptides and in one special case to a larger protein.

Little et al., J. A. Chem. Soc. <u>116</u>: 4893 (1994) disclosed a sequencing method for oligonucleotides using electrospray ionization coupled with Fourier-transform mass spectrometry for oligonucleotides. Fragments were generated using nozzle skimmer dissociation to provide sequence information. The technique was found to be applicable to oligonucleotides from 8 bases to as large as 25 bases.

There remains a need for new sequencing approaches that address the special sequence determination problems presented by synthetic oligonucleotides. Ideally, such a new approach should be rapid and universal, should act directly on the nucleic acid analyte, and should be applicable to any chemically modified oligonucleotides.

### **BRIEF SUMMARY OF THE INVENTION**

The invention provides a universal analytical method for determining the nucleotide sequence of nucleic acid analytes, including any chemically modified oligonucleotides. This new method utilizes electrospray ionization-Fourier transform ion cyclotron resonance mass spectrometry (abbreviated as ESI-FT-ICRMS or ESI-FTMS). This method is extremely rapid and acts directly on the oligonucleotide. It is effective for a variety of nucleic acid analytes, in particular for any chemically modified oligonucleotides which have not previously been successfully sequenced.

The invention provides a method for determining the nucleotide sequence of nucleic acid analytes comprising providing a suitable target nucleotide sequence for ionization by electrospray ionization; exciting the ions of the target; fragmenting the target; and determining the nucleotide sequence of the target by measuring the mass of the resultant fragments. The target ions can be excited and fragmented by any of the known techniques. In one particularly preferred embodiment the target ions are excited by sustained resonance excitation (SORI) and subsequently fragmented by collisionally activated dissociation (CAD). Preferably, the collisionally activated dissociation is by a neutral gas. In another preferred embodiment, the target ions are excited and fragmented by nozzle spray dissociation (NS).

4

### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a schematic drawing of the BioApex FTMS Vacuum System which FTMS system was used in the method according to the invention.

Figure 2 shows an ESI-FTMS mass spectrum of a 25-mer phosphorothioate oligodeoxynucleotide having the sequence 5'-

CTCTCGCACCCATCTCTCTCTCT-3'. The mono-oxygenated 25-mer is also detected, as shown in the inset.

Figure 3, panels A-E, show the ESI-FTMS spectra of a 18-mer phosphorothioate oligodeoxynucleotide having the sequence 5'-AAAAAAAAAAAAAAAAAAAT-3' obtained by SORI CAD. As shown, the entire sequence can be determined by tandem Mass Spectrometry (MS/MS). In panels A-E the following ion series have been identified, respectively, (A) a-B ions; (B) singly charged w ions; (C) doubly charged w ions; (D) triply and quadruply charged w ions; (E) M-n B ions.

Figure 4 shows the ESI-FTMS spectra of an 18-mer phosphorothioate oligodeoxynucleotide having the sequence 5'-AAAAAAAAAAAAAAAAAT-3' obtained by nozzle spray dissociation. Only a partial sequence was obtained.

### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The invention relates to the determination of nucleotide sequences for nucleic acids and their analogs. The patents and publications cited herein are known to those skilled in this field and are hereby incorporated by reference in their entirety.

The invention provides an analytical method for determining the nucleotide sequence of nucleic acid analytes, including chemically modified oligonucleotides. This new method utilizes electrospray ionization-Fourier transform ion cyclotron resonance mass spectrometry. This method is extremely rapid and acts directly on the nucleic acid analyte. It is effective for a variety of nucleic acid analytes, including any chemically modified oligonucleotides which have not previously been successfully sequenced.

The invention provides a method for determining the nucleotide sequence of nucleic acid analytes comprising providing a suitable target nucleotide sequence for ionization by electrospray ionization; exciting the ions of the target; fragmenting the target; and determining the nucleotide sequence of the target by measuring the mass of the resultant fragments. The target ions can be excited and fragmented by any of the known techniques. In one preferred embodiment the ions are excited by sustained resonance excitation (SORI) and fragmented by collisionally activated dissociation (CAD). Preferably, the collisionally activated dissociation is by a neutral gas. In another preferred embodiment the target ions can be excited and fragmented by nozzle spray dissociation (NS).

The nucleic acid analyte can be a naturally occurring or synthetic polynucleotide or oligonucleotide, including oligonucleotides having chemically modified internucleoside linkages, sugar backbones or nucleoside bases. For purposes of the invention, the term "oligonucleotide" includes polymers of two or more deoxyribonucleoside, ribonucleoside or 2'-O-substituted ribonucleoside monomers, or any combination thereof. Preferably, such oligonucleotides will

have from about 2 to about 100 monomers, and most preferably from about 8 to about 70. Such monomers may be coupled to each other by any of the numerous known internucleoside linkages. In certain preferred embodiments, these internucleoside linkages may be phosphodiester, phosphotriester, phosphorothioate, or phosphoramidate linkages, or combinations thereof. The term oligonucleotide also encompasses such polymers having chemically modified bases or sugars and/or having additional substituents, including without limitation lipophilic groups, intercalating agents, diamines and adamantane. For purposes of the invention the term "2'-O-substituted" means substitution of the 2' position of the pentose moiety with an -O-lower alkyl group containing 1-6 saturated or unsaturated carbon atoms, or with an -O-aryl or allyl group having 2-6 carbon atoms, wherein such alkyl, aryl or allyl group may be unsubstituted or may be substituted, e.g., with halo, hydroxy, trifluoromethyl, cyano, nitro, acyl, acyloxy, alkoxy, carboxyl, carbalkoxyl, or amino groups; or such 2' substitution may be with a hydroxy group (to produce a ribonucleoside), an amino or a halo group, but not with a 2'-H group.

The preferred target is a synthetic polynucleotide or oligonucleotide having chemically modified internucleoside linkages, sugar backbones or nucleoside bases. Preferably, a suitable target will be solubilized in a volatile organic solvent. More preferably the volatile solvent includes an alcohol or acetonitrile. In a particularly preferred embodiment the volatile solvent is an approximately equal ratio mixture of water and organic solvent. Useful alcohols include, without limitation, propanol, isopropanol, methanol and ethanol. The volatile solvent may also include low concentrations of organic or inorganic bases, for example piperidine.

The targets are ionized by ESI using a voltage differential between the needle and the end cap. For example the voltage of the needle can be set at ground and the voltage of the end cap can be set at a positive level. The target will be negatively ionized when sprayed through the needle and will be attracted to the positively charged end cap. Alternatively the voltage of the needle could be

set to a negative level, and the target will be attracted to the relatively positively charged end cap. In addition, the level of ionization can be manipulated by increasing or decreasing the level of pneumatic flow on the needle or by altering the differential voltage between the needle and the end cap.

After the target is ionized by spraying it through the needle towards the end cap, the target must be dried. Drying preferably is done at a temperature of at least 150°C, and more preferably drying is done at a temperature between 150-250°C. Drying also can be in the presence of a gas, for example a neutral gas such as carbon dioxide.

The ESI ionized targets are then transmitted through a capillary, a tube lens, and a beam skimmer, guided by a series of three quadrupoles through five stages of differential pumping to an open cylindrical ion cyclotron resonance (ICR) cell. The ionized target is excited, fragmented and the nucleotide sequence is determined.

When the nucleotide sequence is to be determined by CAD the molecular weight of the intact target molecule first can be determined by mass spectrometry. The ionized target is then excited and fragmented. For example, excitation can be by SORI. When using SORI, the level of ion excitation can be manipulated by the frequency shift of the excitation and the bursting time of excitation. Preferably the ions are excited by single shot excitation. The frequency shift is preferably from 0 to 2 kHz, and the attenuation should be adjusted to produce a pulse to pulse range

(p-p) in the range of 0 to 10 V. The bursting time is preferably from 0 to 1000 ms, more preferably 100 to 400 ms. In one particularly preferred embodiment the target ions are excited at a frequency shift of ±650 Hz with attenuation of 35 dB (2V p-p) and a bursting time of 300 ms. After SORI excitation the targets are fragmented by CAD. Prior to CAD all but the selected ions can be ejected from the cell. The excited targets are collided with a neutral gas to fragment the molecule. The extent of collisions are controlled to limit the extent of target

fragmentation. In one particularly preferred embodiment the collision gas is carbon dioxide. The molecular weight of the resultant fragments are then determined by mass spectrometry and the nucleotide sequence is determined.

Excitation and fragmentation also can be done by NS dissociation. For NS dissociation the ionized target is dissociated while it is guided from through the capillary to the ICR cell. For NS dissociation the voltage differential between the capillary exit and the skimmer is preferably between ±0 to ±500 V. In particularly preferred embodiments the voltage differential is 140 V.

The following examples are intended to further illustrate certain preferred embodiments of the invention and are not intended to limit the scope of the invention.

# Example 1 Preparation of an appropriate target

### Example 2

### Sequence determination of a nucleic acid analyte using ESI-FTMS

The oligodeoxynucleotide target prepared according to Example 1 was solubilized in 100:100:1 water:isopropoanol:piperidine and sprayed through a needle towards the end cap in the negative ion mode with pneumatic assistance. The sequencing was performed in an ESI-FTMS, the 7.0 tesla Bruker Apex 70E system (Bruker Analytical Systems, Inc., Billerica, MA) a schematic of which is shown in Figure 1. The pressure in the source chamber and the analyzer chamber was  $4 \times 10^{-6}$  torr and  $7 \times 10^{-10}$  torr, respectively. Carbon dioxide was used both for drying and sheath flow. The drying gas temperature was at  $200^{\circ}$ C.

To sequence the oligonucleotide using the SORI CAD method, the precursor ions were activated by single shot excitation with frequency shift of ±650 Hz, attenuation of 35 dB (2V p-p), and bursting time of 300 ms. Carbon dioxide was used as the collision gas. The voltages were as follows: needle, ground; capillary entrance, 4 kV, end cap, 3.5 kV, cylinder, 3.5 kV; skimmer,

-5 V; capillary exit, -80 V. The analyzed mass spectra are shown in Figures 3A-E.

To sequence the oligodeoxynucleotide using nozzle skimmer dissociation, the voltages were as follows: needle, ground; capillary entrance, 4 kV, end cap, 3.5 kV, cylinder, 3.5 kV; and the voltage differential between the capillary exit and the skimmer was 140 V. The analyzed mass spectra is shown in Figure 4.

### What is claimed is:

1. A method for determining the nucleotide sequence of a nucleic acid analyte having at least one modified internucleoside linkage, sugar backbone or nucleoside base, the method comprising:

- a) providing a suitable target nucleotide sequence for ionization by electrospray ionization, such target comprising a nucleic acid analyte having at least one modified nucleic acid;
- b) exciting the ions of the target;
- c) fragmenting the target; and
- d) determining the nucleotide sequence of the target.
- 2. The method of claim 1 wherein step b further comprises exciting the ions by sustained off resonance excitation.
- 3. The method of claim 2 wherein said excitation is a single shot excitation.
- 4. The method of claim 1 wherein step c further comprises fragmenting the target by collisionally activated dissociation.
- 5. The method of claim 4 wherein said collisionally activated dissociation is by a neutral gas.
- 6. The method of claim 5 wherein said neutral gas is carbon dioxide.
- 7. The method of claim 1 wherein steps b and c are nozzle skimmer dissociation.

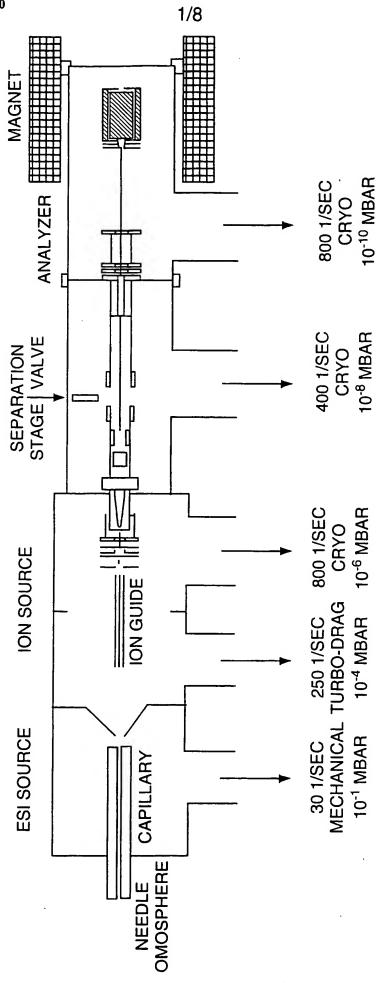
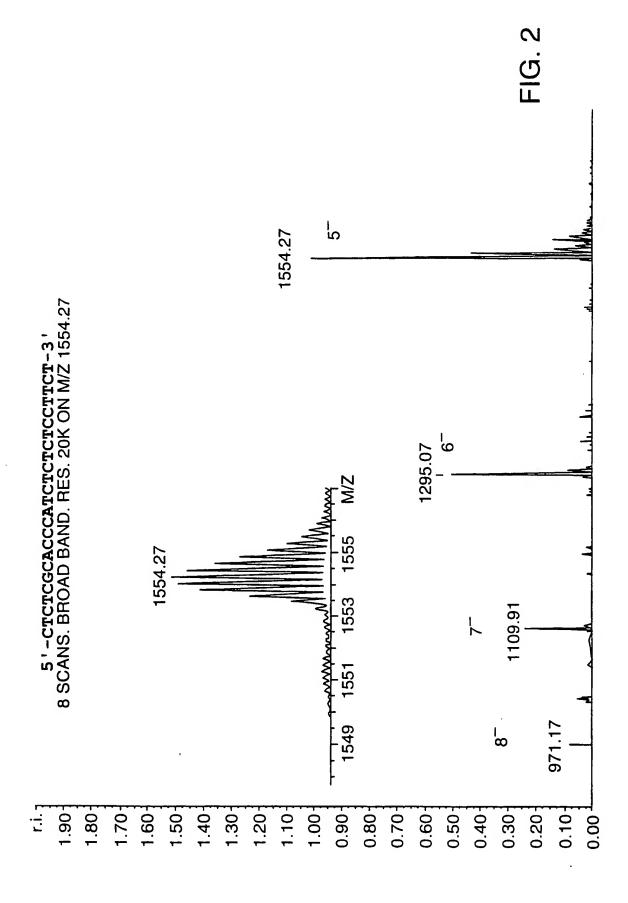
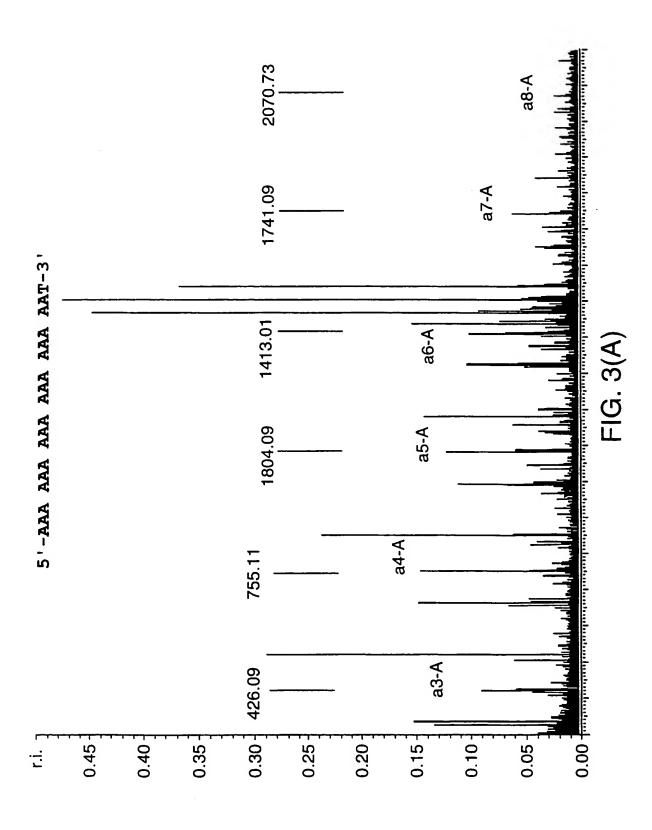
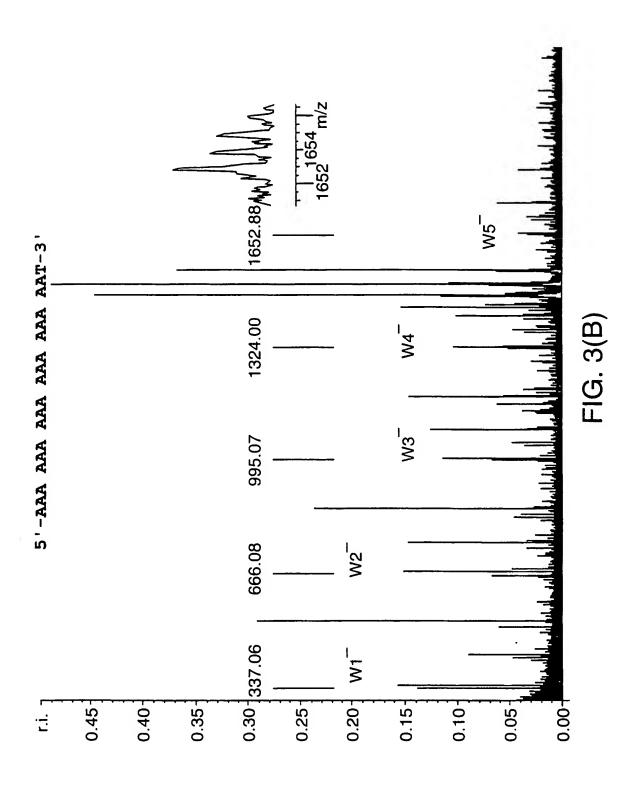
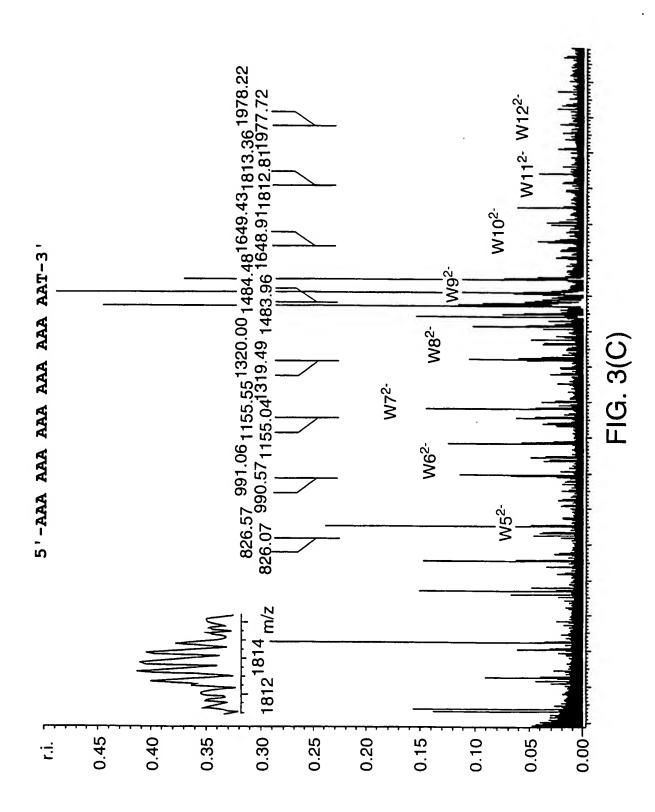


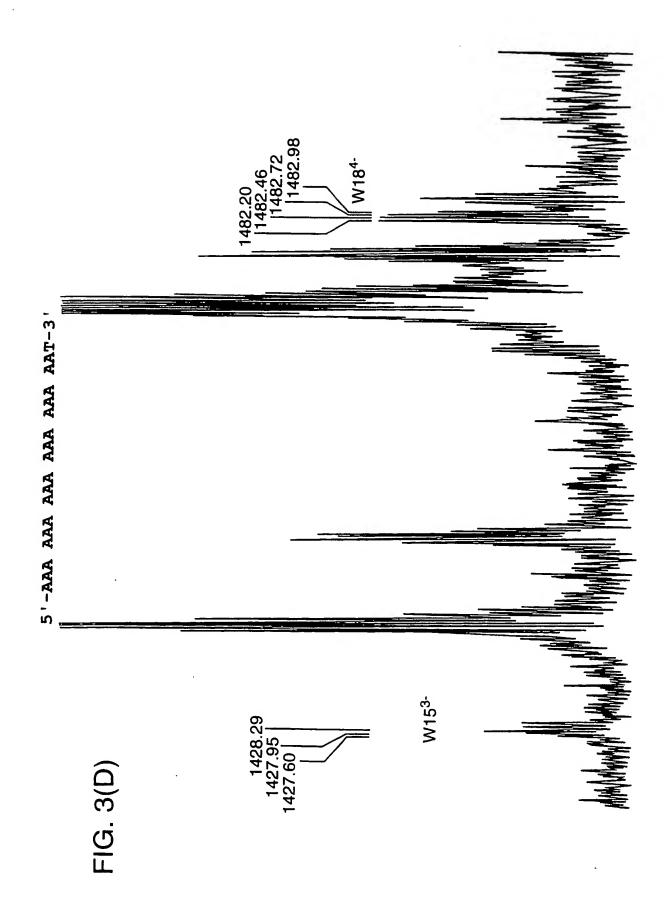
FIG. 1

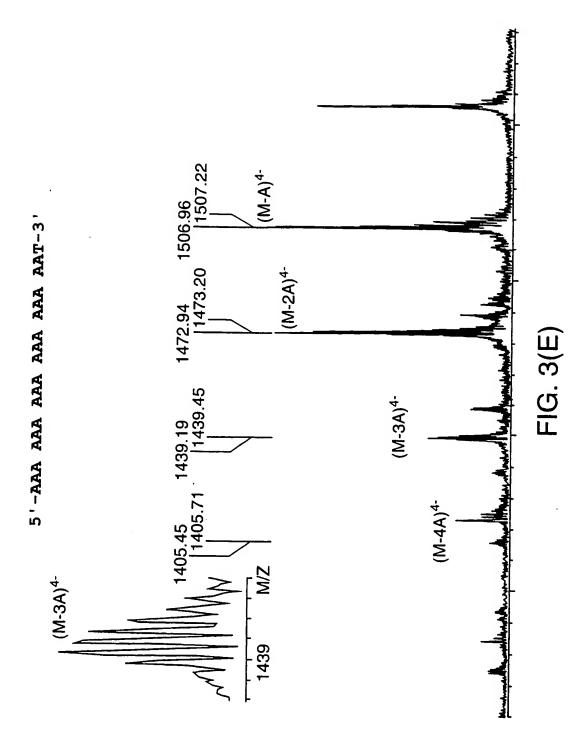


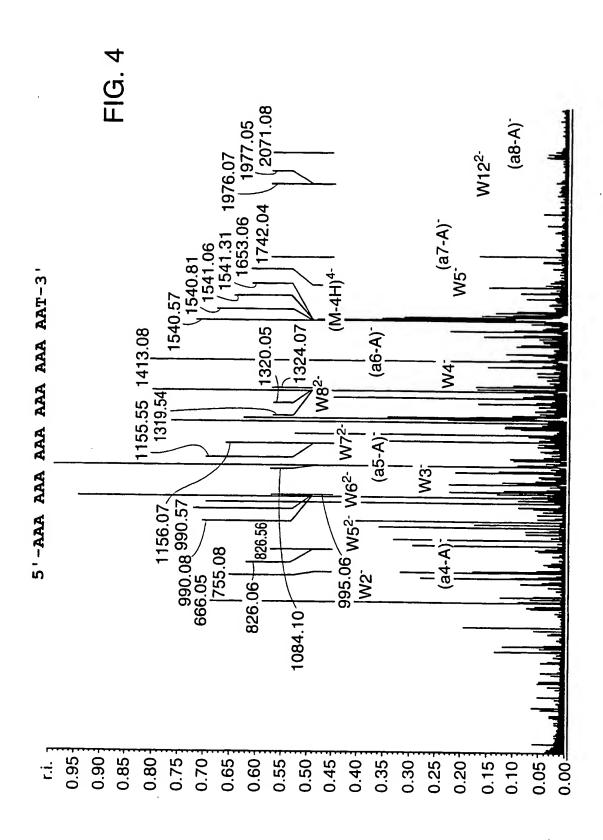












## INTERNATIONAL SEARCH REPORT

Inti tional Application No PCT/US 98/04919

laim No.		
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combined with one or more other such documents, such combination being obvious to a person skilled in the art.  "&" document member of the same patent family  Date of mailing of the international search report		
•		

### INTERNATIONAL SEARCH REPORT

Intraction No PCT/US 98/04919

		PCT/US 98/04919
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category '	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 29431 A (SEQUENOM INC) 26 September 1996 see abstract see page 2, line 19 - page 3, line 16 see page 10, line 19 - line 22 see page 16, line 23 - line 28 see page 20, line 16 - line 23; claims 1,8,10,12-16,19,21; figure 12	1,4
Α	TALBO G ET AL: "ASPECTS OF THE SEQUENCING OF CARBOHYDRATES AND OLIGONUCLEOTIDES BY MATRIX—ASSISTED LASER DESORPTION/IONIZATION POST-SOURCE DECAY" RAPID COMMUNICATIONS IN MASS SPECTROMETRY, vol. 10, no. 1, 1996, pages 100-103, XP002047881 see abstract	·
A	WO 95 20680 A (HYBRIDON INC) 3 August 1995 see the whole document	
P,Y	WANG B H ET AL: "Sequencing of modified oligonucleotides using in-source fragmentation and delayed pulsed ion extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry" INTERNATIONAL JOURNAL OF MASS SPECTROMETRY AND ION PROCESSES, vol. 169-170, December 1997, page 331-350 XP004109255 see the whole document	
P,Y	WO 98 03684 A (HYBRIDON INC) 29 January 1998 see abstract see page 1, line 1 - page 3, line 25; claims 1-3; figure 12	1

### INTERNATIONAL SEARCH REPORT

information on patent family members

Int tional Application No PCT/US 98/04919

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9632504	Α	A 17-10-1996	AU	5544696 A	30-10-1996
			EP	0830460 A	25-03-1998
WO 9629431	Α	26-09-1996	US	5605798 A	25-02-1997
			AU	5365196 A	08-10-1996
			CA	2214359 A	26-09-1996
			EP	0815261 A	07-01-1998
WO 9520680	A	03-08-1995	US	5525470 A	11-06-1996
			AT	153078 T	15-05-1997
			AU	1734695 A	15-08-1995
			CA	2180722 A	03-08-1995
			DE	69500303 D	19-06-1997
			DE	69500303 T	30-10-1997
			EP	0741799 A	13-11-1996
			JP	9508282 T	26-08-1997
WO 9803684	Α	29-01-1998	AU	4042597 A	10-02-1998